



THE UNIVERSITY *of* EDINBURGH

Edinburgh Research Explorer

The potential of biobanked liquid based cytology samples for cervical cancer screening using Raman spectroscopy

Citation for published version:

Traynor, D, Duraipandian, S, Bhatia, R, Cuschieri, K, Martin, CM, O'Leary, JJ & Lyng, FM 2019, 'The potential of biobanked liquid based cytology samples for cervical cancer screening using Raman spectroscopy', *Journal of biophotonics*. <https://doi.org/10.1002/jbio.201800377>

Digital Object Identifier (DOI):

[10.1002/jbio.201800377](https://doi.org/10.1002/jbio.201800377)

Link:

[Link to publication record in Edinburgh Research Explorer](#)

Document Version:

Peer reviewed version

Published In:

Journal of biophotonics

General rights

Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy

The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.



Title: The potential of biobanked liquid based cytology samples for cervical cancer screening using Raman spectroscopy.

Damien Traynor^{*1}, Shiyamala Duraipandian¹, Ramya Bhatia², Kate Cuschieri², Cara M. Martin^{3,4}, John J. O’Leary^{3,4}, Fiona M. Lyng¹

¹DIT Centre for Radiation and Environmental Science, Focas Research Institute, Dublin Institute of Technology (DIT), Kevin St, Dublin 8, Ireland.

² HPV Research Group, Division of Pathology, University of Edinburgh, Queens Medical Research Institute, 47 Little France Crescent, Edinburgh EH16 4TJ

³ Discipline of Histopathology, University of Dublin Trinity College, Dublin, Ireland

⁴ Emer Casey Molecular Pathology Research Laboratory, The Coombe Women and Infants University Hospital

*Author to whom correspondence should be addressed;

Damien Traynor

DIT Centre for Radiation and Environmental Science

Focas Research Institute

Dublin Institute of Technology

Kevin St

Dublin 8

Ireland

E mail Damien.traynor@dit.ie,

Abstract

Patient samples are unique and often irreplaceable. This allows biobanks to be a valuable source of material. The aim of this study was to assess the ability of Raman spectroscopy to screen for histologically confirmed cases of Cervical Intraepithelial neoplasia (CIN) using biobanked liquid based cytology (LBC) samples. Two temperatures for long term storage were assessed; 80°C and -25°C. The utility of Raman spectroscopy for the detection of CIN was compared for fresh LBC samples and biobanked LBC samples. Two groups of samples were used for the study with one group associated with disease (CIN3) and the other associated with no disease (cytology negative). The data indicates that samples stored at -80°C are not suitable for assessment by Raman spectroscopy due to a lack of cellular material and the presence of cellular debris. However, the technology can be applied to fresh LBC samples and those stored at -25°C and is, moreover, effective in the discrimination of negative samples from those where CIN 3 has been confirmed. Pooled fresh and biobanked samples are also amenable to the technology and achieve a similar sensitivity and specificity for CIN 3. This study demonstrates that cervical cytology samples stored within biobanks at temperatures that preclude cell lysis can act as a useful resource for Raman spectroscopy and will facilitate research and translational studies in this area

Introduction

Every year millions of cervical Pap tests are performed throughout the world in countries for purposes of cervical screening. Most Pap tests are performed through use of liquid based cytology (LBC) where cervical cells are collected before deposition into a volume of liquid preservative. As not all the material is required for cytological assessment, the surplus, which would ordinarily be discarded, can be stored within tissue biobanks with due process of governance. Biobanks constitute a valuable source of material which may support a number of studies including those on the natural history of disease, evaluation of screening practices, vaccination effectiveness or the development of new technologies to support screening and disease management^[1-3].

There were an estimated 527,600 new cervical cancer cases and 265,700 deaths from cervical cancer worldwide in 2012^[4]. This demonstrates the importance of both cervical screening to reduce the burden of disease and also investment in research into new technologies that can improve the performance and “reach” of cervical screening. Different collection media for liquid based cytology exist however one of the more common media is PreservCyt (Hologic). PreservCyt is a methanol based solution that preserves cell morphology via fixation. Fixation

is routinely employed as it allows a “snapshot” of a cell’s physical and biochemical state to be assessed. Methanol is an organic solvent that preserves cells through dehydration and precipitation of proteins^[5]. Fixation is important given that sample collection and assessment is not performed concurrently. In addition to supporting routine screening, fixation of cells also supports longer term storage of residual material in biobanks.

The advantage of the Pap test is that it is a widely accepted screening based test with a high specificity of 95-98% and a sensitivity of 74-96%^[6]. The variability in the rates of sensitivity are due to sampling technique and the variability of the cytology based screening. This can result in unnecessary gynaecological referral and patient recall. Persistent infection with high risk human papillomavirus (HPV), is accepted as the major cause of cervical pre-cancer and cancer^[7]. HPV DNA testing has a higher sensitivity (>95%) but lower specificity (~ 84%) than the Pap test^[7]. These tests are expensive, time-consuming and provide no information on cervical cytopathology.

Current methods for detection of cervical cancer and pre-cancer (CIN) are limited and there is an unmet clinical need for new screening or diagnostic tests. Recently Raman spectroscopy has shown potential as a tool for screening and diagnosis of cervical lesions and cancer^[8-10]. Raman spectroscopy is based on inelastic light scattering where a sample is illuminated by a monochromatic laser light and interactions between the incident photons and molecules in the sample result in the scattering of the light. The coupling of the light generates vibrations within the sample which are characteristic of the chemical structure. This means that the position, peaks and shape of the Raman bands carry information about the molecular makeup of the sample. The Raman spectrum of cells and tissues is made up of contributions from

many biochemical components including DNA, RNA, proteins, lipids and carbohydrates ^[11]. Raman spectroscopy can offer a label free non-destructive method for cervical cancer screening. It is an objective method, less reliant on operator performance than cytology and potentially more specific than HPV testing.

Due to confounding factors such as sample collection, blood contamination and sample variability, few studies have been performed using Raman spectroscopy on cervical cytology samples and none to our knowledge have investigated the potential of utilising biobanked LBC samples. The aim of this feasibility study was to assess the utility and performance of Raman spectroscopy for the detection of CIN using biobanked LBC samples. Samples stored at -80°C and -25°C were assessed and the ability of Raman spectra to delineate disease from no disease was determined. Additionally, Raman spectroscopy was assessed in un-banked LBC samples as a comparator.

Materials and Methods

Sample collection

Two classes of samples were used for the study, classed as disease and no disease. Samples with no disease were defined as cytology negative and HPV negative whereas samples with disease were those associated with a histologically confirmed CIN3 with a HPV positivity according to HPV DNA and mRNA status. All samples were recruited from patients presenting at a colposcopy clinic for the first time, and had no prior history of disease. Samples were collected from each patient according to the standard operating procedure issued by Cervical Check Irelands national cervical cancer screening programme and the NHS Scottish cervical screening programme. Both procedures are similar and all samples were biobanked using the same methodology.

133 samples were used in total for this study of which 64 were LBC biobanked samples; 32 with no disease (cytology negative) and 32 with disease (CIN 3). Biobanked samples were

provided by the Scottish HPV Archive, a research tissue biobank set up to facilitate HPV associated research.

Ethical approval for use of the samples was obtained from the East of Scotland Research Ethics Service - Tayside committee. Biobanked LBC samples used for this study had been sedimented with the cellular pellet transferred into a 4.5 ml vial for long term storage in PreservCyt. After transit, samples were re-constituted to a volume of 20 ml fresh PreservCyt solution to resemble the original LBC specimen from which the sample was derived.

A further 64 non biobanked “fresh” LBC samples, 32 with no disease (cytology negative) and 32 with disease (CIN 3), were collected in PreservCyt solution from the Coombe Women and Infants University Hospital (CWIUH), Dublin, Ireland, as part of routine cytological screening. Ethical approval for use of anonymised samples for the study was granted by the CWIUH Research Ethics Committee (no. 28-2014). A further 5 fresh LBC samples with disease (CIN 3) were collected and split into two separate vials. One vial from each sample underwent the standard biobanking process and was stored for 3 weeks, while the other was stored at room temperature.

ThinPrep

The samples were then prepared using the ThinPrep 2000 processor (Hologic Inc., Marlborough, MA 01752). The ThinPrep process begins with the patient’s gynaecological sample being collected by the clinician using either a cervical broom or brush. The brush/broom is then rinsed in the specimen vial containing PreservCyt transport medium

(ThinPrep Pap Test; Cytoc Corporation, Boxborough, Mass). The ThinPrep sample vial is then capped, labelled and sent to the lab to be processed. The ThinPrep processor homogenizes the sample by spinning the filter, creating shear forces that breaks up any clumped material (blood, mucin and non-diagnostic material). The cells are then transferred onto a polycarbonate filter membrane of the TransCyt filter and transferred onto a glass slide to produce a circular monolayer of cells approx. 20 mm in diameter. The slide is then ejected into a fixative bath of 95% ethanol

Raman spectroscopy

All Raman analysis was performed using a HORIBA Jobin Yvon XplorRA system (Villeneuve d'Ascq, France), which incorporates an Olympus microscope BX41 equipped with a X100 objective (MPlan, Olympus, NA = 0.9). A 532 nm diode laser source was used. Laser power was set to 100% resulting in 16 mW at the objective. The confocal hole coupled to a slit aperture of 100 μm , was set at 100 μm , for all the measurements. The system was pre-calibrated to the 520.7 cm^{-1} spectral line for silicon. A 1200 lines per mm grating was used. The backscattered light was measured using an air-cooled CCD detector (Andor, 1024x256 pixels). The spectrometer was controlled by Labspec V6.0 software. Two accumulations of 30 seconds were performed on each cell nucleus selected. Raman spectra were acquired from the nuclei of 20 randomly selected morphologically normal superficial and intermediate cells from each unstained Pap smear.

Data pre-processing and analysis

Data was normalised and analysed using Matlab software (Mathworks) and specific scripts developed and adapted for uploading of the spectra and their pre-processing, including

smoothing (Savitzky-Golay $K=5$, $K=13$), baseline correction (Rubberband) and vector normalization. The spectra were also corrected for the glass background using a linear least-squares method with non-negative constraints (NNLS). The least-squares model was developed using spectra from the Thinprep glass slides and selected pure biochemicals (e.g., actin, glycogen, RNA, DNA, etc.) that approximate the biochemical composition of cervical cells. The data was mean centred and subjected to partial least squares discriminant analysis (PLS-DA). PLS-DA involves the creation of latent variables to maximise the co variation between known datasets and the response variable which they are regressed against. PLS-DA is a form of analysis that has the ability to distinguish between known classifications of samples and its aim is to find latent variables and directions to maximise separation in a multivariate space ^[12]. To validate the method, leave one patient out cross validation was performed which involved data from one patient sample being removed from the model, with this process repeated until all patient samples were left out once

Results

-25°C Vs -80°C biobanked LBC samples

The samples stored at -25°C presented with intact cellular morphology (Figure 1(A)) and allowed for high quality spectra to be recorded (Figure 1(B)). The samples stored at -80°C presented with cell lysis, cellular debris and very little cellular material which prevented the recording of spectra (Figure 1(C)). One possible explanation for this, is the freeze thaw effect which is commonly used to lyse bacteria and mammalian cells. Storing cells at -80 °C in PreservCyt without any Dimethyl Sulfoxide and bring up to room temperature can cause the

cells to contract during the thawing process resulting in cell lysis. As a result, only biobanked samples previously stored at -25°C were used for this study.

Negative Vs CIN 3 (fresh LBC samples) Model

In order to determine if biobanked LBC samples could be used to discriminate no disease (cytology negative) from disease (CIN 3) using Raman spectroscopy, fresh (non biobanked) LBC samples were first examined as a control. Figure 2(A) shows mean spectra of Negative Vs CIN 3. Figure 2(B) is a latent variables (LV) scores scatter plot of LV1 and LV2 which shows good discrimination along LV1 and LV2. The loadings shown in Figure 2(C), show that the discrimination is based around Raman peaks at 484 (glycogen), 575 (glycogen), 881 (nucleic acids), 1004 (proteins Phenylalanine), 1139, 1238 (proteins Amide III), 1487 (proteins), 1575 (nucleic acids), 1605 (proteins) and 1669 cm^{-1} (proteins Amide I). The LV2 loadings show discrimination is based around 1238 (proteins), 1381 (glycogen), 1450 (proteins and lipids), 1642 (proteins) and 1669cm^{-1} (proteins) (see table 1) ^[13]. The PLS-DA prediction plot shown in Figure 2(D) and has a sensitivity of 86% and a specificity of 90% for CIN 3.

Negative Vs CIN 3 (Biobanked LBC samples) Model

In order to determine if the biobanked samples can be used in a similar fashion to the fresh samples, negative and CIN3 biobanked samples were compared. Figure 5.3(A) shows the mean spectra of biobanked Negative samples Vs CIN 3. Figure 5.3(B) is a latent variables (LV) scores scatter plot of LV1 and LV2 which shows good discrimination along LV1 and

LV2. The loadings from LV1 are shown in Figure 5.3(C) and show that discrimination is based around Raman peaks, 622 (proteins), 640 (proteins), 775 (proteins), 850 (proteins), 1122 (proteins), 1152 (proteins), 1207 (proteins), 1450 (proteins), 1560 (proteins), 1605 (proteins), 1642 (proteins) and 1669 cm^{-1} (proteins). LV2 loadings show discrimination is based on 1123 (proteins, lipids, carbohydrates), 1338 (proteins) and 1605 cm^{-1} (proteins). Raman peaks assigned to phenylalanine 1004 cm^{-1} show a slight shift between $1003\text{--}1004\text{ cm}^{-1}$ which is most likely attributed to the methanol based fixation method which suggests a change in the conformation of the phenylalanine protein^[14]. The PLS-DA prediction plot shown in Figure 5.3(D) and has a sensitivity of 91% and a specificity of 92% for CIN 3.

Biobanked Vs non-Biobanked samples

5 fresh CIN 3 patient samples were split into two separate vials. One vial from each sample was frozen as described earlier and the other stored at room temperature. Figure 4(A) show the mean spectra for biobanked CIN 3 samples and the same samples kept at room temperature for 3 weeks after collection. The mean spectra appear identical. There does not appear to be a difference between the fresh and biobanked samples. The latent variable scatter scores plot (Figure 4(B)) shows slight discrimination between the sample types which is most likely due to internal sample variability^[15] (LBC samples are variable by nature) and the low number of spectra recorded (60 for room temperature/biobanked). The PLS-DA prediction plot (Figure 4(C)) has a sensitivity of 29% and specificity of 88% for biobanked samples indicating poor discrimination between the two groups.

In order to determine if we could mix fresh and biobanked samples together and still achieve a sensitivity and specificity similar to the fresh and biobanked models, 15 biobanked CIN 3 samples were mixed with 15 fresh CIN 3 samples and compared with 15 negative biobanked/ 15 fresh negative samples. Figure 5(A) shows the latent variable scatter scores plot of the model and we can see clear discrimination between the sample types across LV1 and LV2. The LV1 loadings (Figure 5(B)) show that discrimination is based on 482, (glycogen), 1443 (proteins, lipids) 1487 (proteins), 1605 (proteins) 1669 cm^{-1} (proteins) while LV2 shows the discrimination is based around 486 (glycogen), 851 (proteins), 1152 (proteins), 1381 (glycogen), 1450 (proteins/lipids), 1575 (nucleic acids) and 1669 cm^{-1} (proteins). The loadings show similarities to both the fresh and biobanked loadings, but overall show that glycogen and proteins are the main discriminating factor between negative and CIN 3 samples. PLS-DA prediction plot has a sensitivity of 94% and a specificity of 95% for CIN 3 (Figure 5(C)).

Discussion

From the results it is clear that samples biobanked at -80°C are not suitable for screening using Raman spectroscopy due to a lack of cellular material and the presence of cellular debris.

Spectral differences between fresh negative and CIN 3 samples were observed with regards to glycogen, nucleic acids and proteins. CIN 3 cells often contain little to no glycogen, hence the use of Lugol's solution to visualise abnormal cells in colposcopy^[16]. The discrimination associated with changes in proteins and DNA is consistent with the neoplastic changes that occur in CIN 3 supported by persistent HPV infection such as increased cell cycling with coincident increase in replication and levels of nucleic acids^[15]. The PLS-DA prediction plot gives a sensitivity of 86% and a specificity of 90% for CIN3.

Negative Vs CIN 3 biobanked sample results showed that discrimination was driven solely by proteins. Raman peaks associated with nucleic acids/ DNA are not as strongly present as they are in the non-biobanked samples. Long term storage of biobanked samples is likely to have led at least to an element of nucleic acids degradation which would explain why nucleic acid is not discriminatory between negative and CIN3 samples. However the PLS-DA prediction plot (Figure 3(D)) does show slightly higher sensitivity (91%) and specificity (92%) rates when compared to fresh samples (86% sensitivity and 90% specificity) indicating that biobanking at -20°C does not preclude discrimination of negative and CIN 3 samples on Raman spectroscopy.

The same patient samples that had been split in two with half biobanked and the other half stored at room temperature showed no discrimination between the samples. Hence the 3 week period of biobanking at -20°C had no detrimental effects on the physical or biochemical properties of the samples. The mixed model showed that biobanked and fresh LBC samples could be combined with an improved sensitivity of 94% and specificity of 95%. A limitation of this study is the inability to use biobanked LBC samples stored at -80°C for Raman spectroscopic analysis as most biobanks will have samples stored at -80°C for long term storage^[3] hence the true potential of using biobanks as a source of patient samples could be lost. Further research in this area should involve the use of different biobank specimens (bronchial and thyroid fine needle aspirations) to investigate any detrimental effects the biobanking process may have on cytological specimens^{17 18}.

Conclusion

Raman spectroscopy can effectively discriminate disease free cervical LBC samples from those with disease (CIN 3) and this is possible using biobank cervical LBC samples stored at -25°C . Pooling samples stored at -25°C with fresh samples does not affect the sensitivity and specificity of Raman spectroscopy for the discrimination of disease. This study demonstrates that biobanks of cervical LBC samples are a useful resource for future Raman spectroscopy studies and will facilitate the further assessment of this technology which shows highly encouraging performance for the detection of cervical disease

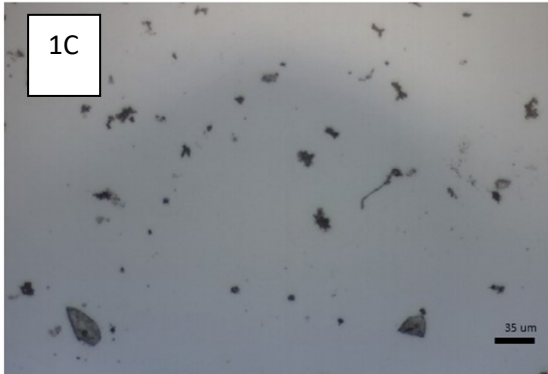
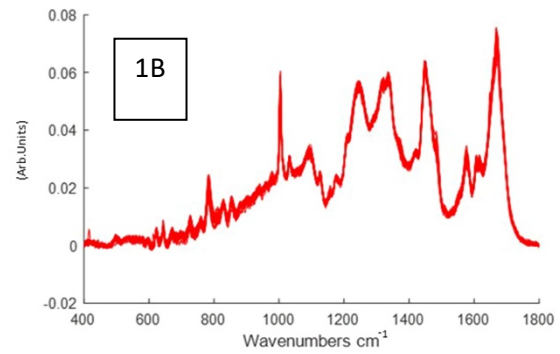
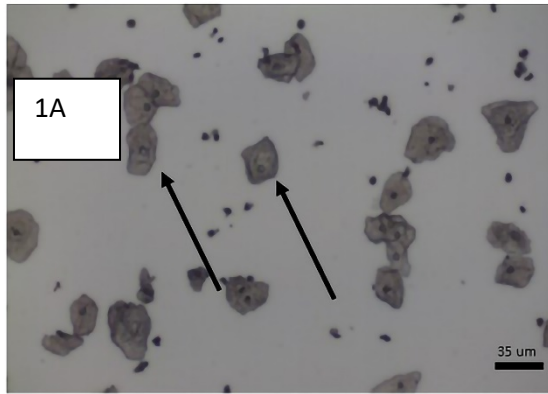


Figure 1 (A) LBC samples stored at -25°C present with intact cellular morphology. Note the presence of superficial and intermediate cells on the unstained slide which were selected for Raman spectral recording. (B) High quality spectra recorded from morphologically normal intermediate and superficial cells in the spectral range $400\text{-}1800\text{cm}^{-1}$. (C) LBC samples stored at -80°C . Note lack of cellular material and presence of cellular debris.

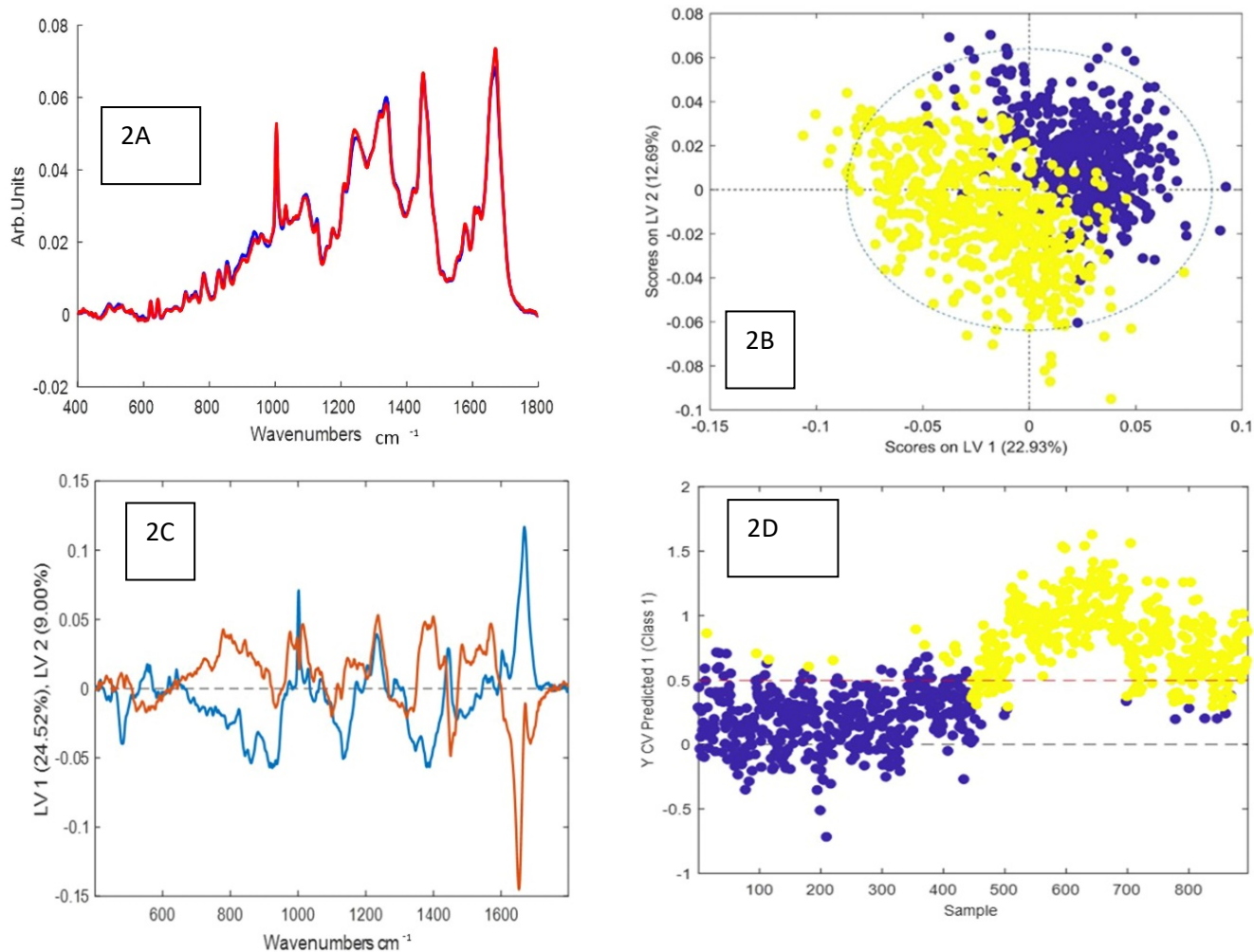


Figure 2 (A) mean spectra of fresh Negative (red) Vs CIN 3 (blue). (B) is a latent variables (LV) scores scatter plot of LV1 and LV2, TN (yellow) Vs CIN 3 (blue). (C) LV1(blue) and LV2 (orange) loadings (D) PLS_DA prediction plot CIN 3 (blue), negative (yellow)

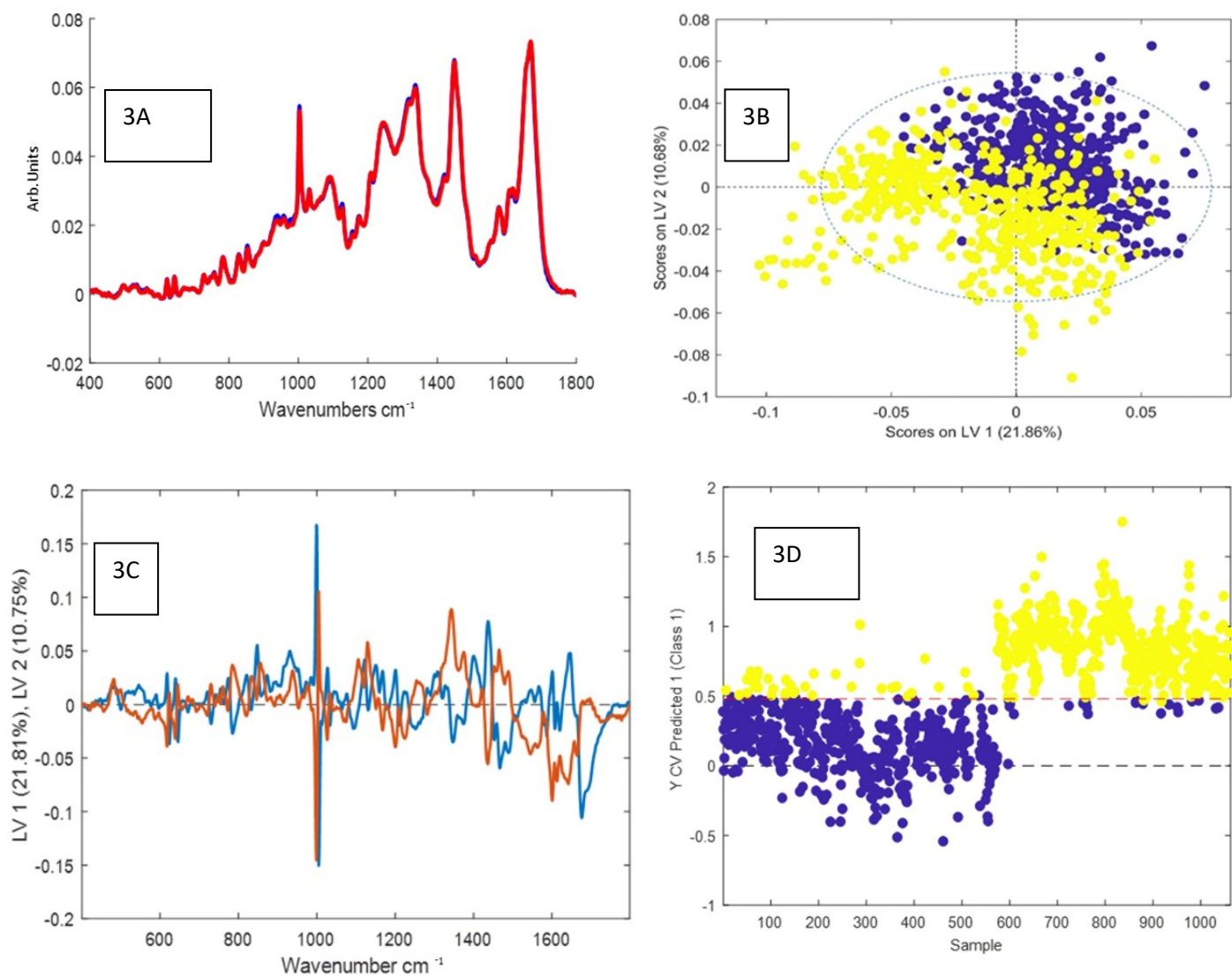


Figure 3 (A) mean spectra of biobanked Negative (red) Vs CIN 3 (blue). (B) latent variables (LV) scores scatter plot of LV1 and LV2, TN (yellow) Vs CIN 3 (blue). (C) LV1 (blue) LV2 (orange). (D) PLS_da prediction plot CIN 3 (blue), negative (yellow).

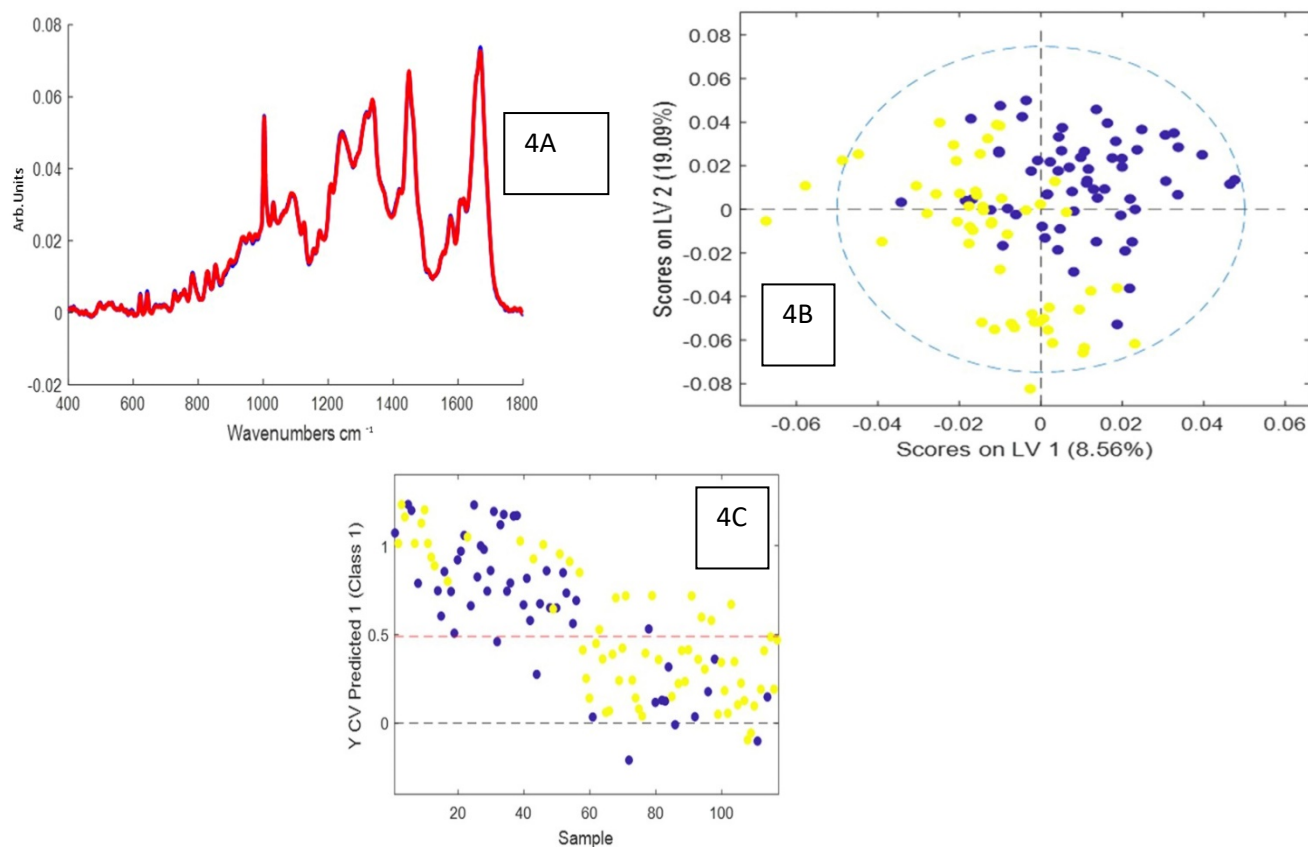


Figure 4 (A) mean spectra of fresh CIN 3 (blue) vs biobanked CIN 3 (red). (B) latent variables (LV) scores scatter plot of LV1 and LV2, fresh CIN 3 (yellow) Vs biobanked CIN 3. (C) PLS-DA prediction plot biobanked CIN 3 (blue) vs fresh CIN 3 (yellow).

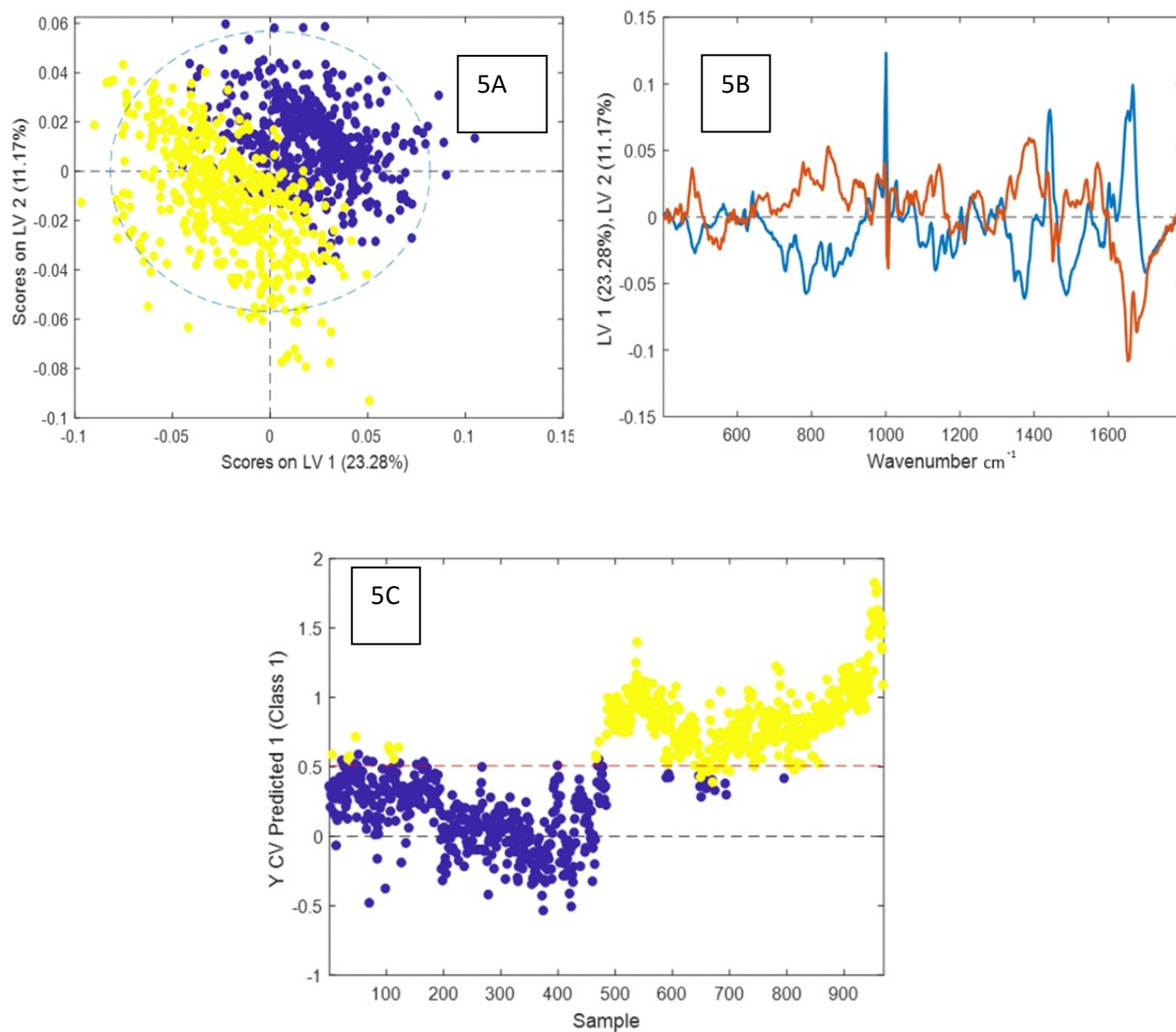


Figure 5 (A) latent variables (LV) scores scatter plot of LV1 and LV2, TN (yellow) Vs CIN 3 (blue). (B) LV1 loadings (blue) and LV2 loadings (orange). (C) PLS-DA prediction plot CIN 3 (blue), negative (yellow)

References

- [1] Fox, J. M. *et al.* (2017) 'Methodology for reliable and reproducible cryopreservation of human cervical tissue', *Cryobiology*. Elsevier Ltd, 77, pp. 14–18. doi: 10.1016/j.cryobiol.2017.06.004.
- [2] Marquez-Curtis, L. A., McGann, L. E. and Elliott, J. A. W. (2017) 'Expansion and cryopreservation of porcine and human corneal endothelial cells', *Cryobiology*. Elsevier Ltd, 77, pp. 1–13. doi: 10.1016/j.cryobiol.2017.04.012.
- [3] Peakman, T. and Elliott, P. (2010) 'Current standards for the storage of human samples in biobanks', *Genome Medicine*, 2(10), pp. 2–4. doi: 10.1186/gm193
- [4] Torre, L. A. *et al.* (2015) 'Global Cancer Statistics, 2012', *CA: a cancer journal of clinicians.*, 65(2), pp. 87–108. doi: 10.3322/caac.21262.
- [5] Troiano, N. W., Ciovacco, W. A. and Kacena, M. A. (2009) 'The Effects of Fixation and Dehydration on the Histological Quality of Undecalcified Murine Bone Specimens Embedded in Methylmethacrylate', *Journal of Histotechnology*, 32(1), pp. 27–31. doi: 10.1179/his.2009.32.1.27.
- [6] Kitchener, H. C. *et al.* (2011) 'Automation-assisted versus manual reading of cervical cytology (MAVARIC): A randomised controlled trial', *The Lancet Oncology*. Elsevier Ltd, 12(1), pp. 56–64. doi: 10.1016/S1470-2045(10)70264-3.
- [7] Ronco, G. *et al.* (2010) 'Efficacy of human papillomavirus testing for the detection of invasive cervical cancers and cervical intraepithelial neoplasia: a randomised controlled trial',

The Lancet Oncology, 11(3), pp. 249–257. doi: 10.1016/S1470-2045(09)70360-2.

[8] Lyng, F. Traynor, D. (2015) 'Raman spectroscopy for cytopathology of exfoliated cervical cells', *Analytical and Bioanalytical Chemistry*, 407(27), pp. 8279–8289.

[9] Ramos, I. R. *et al.* (2016) 'Raman spectroscopy for cytopathology of exfoliated cervical cells', in *Faraday Discuss.*, pp. 187–198. doi: 10.1039/C5FD00197H.

[10] Rubina, S. C, Murali, K. (2015) 'Raman spectroscopy in cervical cancers: An update', *Journal of cancer research and therapeutics*, 11(1), pp. 10–17.

[11] Santos, I. P, Barroso, E. M. (2017) 'Raman spectroscopy for cancer detection and cancer surgery guidance: translation to the clinics', *Analyst*, 142, pp. 3025–3047.

[12] Brereton, R. G., Lloyd, G. R. (2014) 'Partial least squares discriminant analysis: Taking the magic away', *Journal of Chemometrics*, 28(4), pp. 213–225. doi: 10.1002/cem.2609.

[13] Movasaghi, Z., Rehman, S. and U., R. I. (2007) 'Raman spectroscopy of biological tissues.', *Applied Spectroscopy Reviews*, 42, pp. 493–541.

[14] Meade, A. (2010) 'Studies of Chemical fixation effects in Human cell lines using Raman Microspectroscopy', *Bioanalytical chemistry*, 369(5), pp. 1781–1791. doi: 10.1007/s00216-009-3411-7

[15] Serafin-Higuera, I. *et al.* (2016) 'Differential proteins among normal cervix cells and cervical cancer cells with HPV-16 infection, through mass spectrometry-based Proteomics (2D-DIGE) in women from Southern Mexico', *Proteome Science*. Proteome Science, 14(1), pp. 1–9. doi: 10.1186/s12953-016-0099-4.

[16] CervicalCheck (2013) *Organisational and Clinical Guidance for Quality Assured Colposcopy Services*.

[17] Al-Abbadi MA. Basics of cytology. *Avicenna J Med*. 2011;1(1):18-28.

[18] O'Dea, D. *et al.* (2018) 'Raman spectroscopy for the preoperative diagnosis of thyroid cancer and its subtypes: An in vitro proof-of-concept study', *Cytopathology*, (March). doi: 10.1111/cyt.12636.